基于 MFCC 和 GMM 的昆虫声音自动识别

竺乐庆[1,*]，张　真[2]

（1. 浙江工商大学计算机与信息工程学院，杭州 310018；
2. 中国林业科学研究院森林生态与保护研究所，国家林业局森林保护重点实验室，北京 100091）

摘要：昆虫的运动、取食、鸣叫都会发出声音，这些声音存在种内相似性和种间差异性，因此可用来识别昆虫的种类。基于昆虫声音的昆虫种类自动检测技术对协助农业和林业从业人员方便地识别昆虫种类非常有意义。本研究采用了语音识别领域里的声音参数化技术来实现昆虫的声音自动鉴别。声音样本经预处理后，提取梅尔倒谱系数(Mel-frequency cepstrum coefficient，MFCC)作为特征，并用这些样本提取的 MFCC 特征集训练混合高斯模型(Gaussian mixture model，GMM)。最后用训练所得到的 GMM 对未知类别的昆虫声音样本进行分类。该方法在包含 58 种昆虫声音的样本库中进行了评估，取得了较高的识别正确率(平均精度为 98.95%)和较理想的时间性能。该测试结果证明了基于 MFCC 和 GMM 的语音参数化技术可以用来有效地识别昆虫种类。

关键词：昆虫；种类鉴定；声音处理；自动识别；梅尔倒谱系数；混合高斯模型
中图分类号：Q967　　**文献标识码**：A　　**文章编号**：0454-6296(2012)04-0466-06

Automatic recognition of insect sounds using MFCC and GMM

ZHU Le-Qing[1], ZHANG Zhen[2] (1. College of Computer Science and Information Engineering, Zhejiang Gongshang University, Hangzhou 310018, China; 2. Key Laboratory of Forest Protection of State Forestry Administration, Research Institute of Forest Ecology, Environment and Protection, Chinese Academy of Forestry, Beijing 100091, China)

Abstract: Insects produce various sounds when they are moving, feeding or calling. These sounds exhibit intraspecies similarity and interspecies differences, thus they can be used to discriminate species identities of insects. Automatic detection of insect species through sounds produced by the insects would be very meaningful in giving farm workers or forestry workers a convenient way to recognize insects. In this study we employed the sound parameterization techniques that are frequently used in the field of human speech recognition. Mel-frequency cepstrum coefficients (MFCCs) were extracted from the sound samples after preprocessing, and Gaussian mixture model (GMM) was trained with these MFCC features. Finally, the unknown insect sound samples were classified by the GMM. The proposed method was evaluated in a database with acoustic samples of 58 different insect sounds. The method performed well in terms of both recognition rate and time performance. The average recognition accuracy was as high as 98.95%. The test results proved that sound parameterization techniques based on MFCC and GMM could be used to recognize insect species efficiently.

Key words: Insects; species identification; sound processing; automatic recognition; Mel-frequency cepstrum coefficient (MFCC); Gaussian mixture model (GMM)

对有害昆虫的鉴定和检测的传统方法通常是用诱捕法手工实施，然而，近年来在信号处理和计算机技术上的发展使得通过图像分析和声学检测等方法进行物种的自动鉴定成为可能。由于人工统计物种数量在实施过程中存在困难，用生物声学方法实现物种的自动鉴定的研究越来越广泛。研究人员已把语音识别技术应用到鸟类的分类研究中，如台北科技大学的薛宇志(2010)使用二阶段法识别架构，先判断鸟鸣声的鸣叫声模式，再根据该模式判断鸟的种类。中国科学院动物研究所的科研人员利用语音识别方法，在 4 种雀形目鸟类中进行了种内个体识别的研究(Cheng *et al.*, 2010)。这些研究表明，

基金项目：浙江省自然科学基金项目(Y12F020130)；浙江省教育厅资助项目(Y201119748)；浙江省科技厅资助项目(2010C31108)
作者简介：竺乐庆，女，1972 年 7 月生，浙江嵊州人，博士，副教授，研究方向为图像处理及模式识别，E-mail：zhuleqing@zjgsu.edu.cn
* 通讯作者 Corresponding author, E-mail：zhuleqing@zjgsu.edu.cn
收稿日期 Received：2012-01-16；接受日期 Accepted：2012-03-29

声音在物种分类研究中存在广阔的应用前景。尽管在人类语言识别领域投入的研究已较多，但自动声学物种鉴别仍被认为是模式识别的边缘领域，这方面研究的文献也相对较少。概言之，昆虫的声学鉴定是基于昆虫具有产生声音的能力，这种声音要么是出于沟通需要有意发出的，要么是作为取食、飞行或运动的副产品而发出。如果这种昆虫发出生物声学信号遵循种类一致的声学模式，那么它就可以被用于种类的检测和鉴定。Riede（1998）提出昆虫所发出的声音提供了可靠的分类学线索，因此可被用于生物多样性度量。

昆虫的声学鉴定问题主要分为两个阶段：特征提取和分类识别。所提取的特征在其空间中应有能力区分不同种类的昆虫，而分类器则能在给定特征空间实现这种区分。韩萍（2003）使用频谱分析法和 BP 神经网络识别储粮害虫，并在包含 3 类害虫［米象 *Sitophilus oryzae*（Linne）、玉米象 *Sitophilus zeamais*（Motschulsky）和赤拟谷盗 *Tribolium castaneum*（Herbst）］的库中进行了测试，并取得了 81% 的识别正确率。Chesmore（2001）用时域信号处理和人工神经网络对自动识别直翅目昆虫蚱蜢（grasshoppers）和蟋蟀（crickets）的技术进行了研究，选择了 25 类英国直翅目昆虫作为测试集，初步的结果表明可以获得较高的识别率。Pinhas 等（2008）则开发了一种使用矢量量化（vector quantization，VQ）和混合高斯模型（Gaussian mixture model，GMM）的数学方法实现对红棕象虫（red palm weevil）声学行为的自动检测。Ganchev 等（2007）使用主谐波、脉冲周期和宽度以及 23 个线性倒谱系统（linear frequency cepstrum coefficients，LFCCs），经归一化后作为特征向量，并用概率神经网络（probabilistic neural network，PNN）、GMM、隐马尔可夫模型（hidden Markov model，HMM）等作为分类器识别不同的昆虫声音，提出的方法在北美鸣虫数据集中进行了评估并取得了较高的识别精度。

本研究使用梅尔倒谱系数（Mel-frequency cepstrum coefficient，MFCC）作为从昆虫声音样本中提取的特征，并用混合高斯模型来分类不同的声音。本研究方法在储粮害虫的运动和取食声、土壤昆虫的运动和取食声（Arbogast *et al*.，2000；Brandhorst-Hubbard *et al*.，2001；Zhang *et al*.，2003；Mankin *et al*.，2008a）、土壤昆虫的防御性摩擦声（Vulinec，2000）、树林昆虫的运动和取食声（Thoms，2000；Mankin *et al*.，2008b）、植物昆虫的运动和取食声、翅和腹部的振动声（Hay-Roe and Mankin，2004）等样本库中进行了评估，对识别特定的昆虫种类取得了 98% 以上的识别精度。

1　声音样本来源和预处理

1.1　声音录制及数据获取

本研究所使用的实验材料为昆虫声音库（Mankin，2009），该库由美国农业部农业研究服务中心 Richard Mankin 的研究小组创建。该库中所录制的昆虫声音持续时间介于 3 ~ 60 s 之间。为了使 GMM 分类器有充分的数据用于训练，昆虫声音片段较长的样本用于训练，而短样本片段用于测试以提高识别的速度。对于识别过程而言，一个包含活跃信号的 1.2 s 长的声音片段已足以提取用于识别的有用参数。库中数据被分为两个数据集：一个训练集和一个测试集，训练集中的样本长度要远长于测试集。

1.2　预处理

假设所有的输入昆虫声音是经过采样和量化的数字信号，那么预处理主要包括信号归一化、预加重和分段。

1.2.1　归一化：声音信号的归一化只是把每一个采样值除以本段信号的幅度峰值，即：

$$\hat{X}(i) = x(i) / \max_{0 \leqslant i < n-1} x(i) \tag{1}$$

其中 $x(i)$ 是原始信号，$\hat{x}(i)$ 是归一化后的信号，n 是信号长度。

1.2.2　预加重：由于声音信号随频率的增长呈指数级的衰减，大多能量集中在低频带，而高频成分的信噪比则衰减到一种无法接受的水平。预加重是一种提升高频成分的方法，而低频成分则保持原来的水平，使信号的频谱变得平坦，以便于进行频谱分析或声道参数分析。预加重因子 α 计算如下：

$$\alpha = \exp(-2\pi F \Delta t) \tag{2}$$

其中 Δt 为声音信号的采样周期，则预加重滤波器的 z 传递函数为：

$$H(z) = 1 - \alpha z^{-1} \tag{3}$$

1.2.3　分段：只对测试集的数据集进行了分段操作。较长的声音样本被均匀地分割为长度约 1.2 s 的短样本。

2　MFCC 特征提取

梅尔倒频谱（Mel-frequency cepstrum，MFC）是对声音的短时能量谱的一种表示，它是对非线性梅

尔频标中的对数功率谱进行线性余弦变换所得。梅尔倒谱系数(MFCCs)是共同构成 MFC 的全体系数(Mermelstein, 1976; Davis and Mermelstein, 1980)。它们起源于一种音频片段的倒谱表示。倒谱和梅尔倒谱之间的区别在于，在梅尔频标上均匀分布的频带，比常规倒谱中线性分布的频带，更接近于人类听觉系统的响应。MFCCs 已被广泛应用在语音处理的说话人识别领域(Wu and Cao, 2005; 张万里和刘桥, 2005)，竺乐庆等(2010)对 MFCCs 的特性和计算过程作了较详细的介绍，因此本文只作简略描述，有兴趣的读者可参阅相关文献。

计算 MFCCs 的过程如下：

1)预加重，详见 1.2.2。

2) 加 Hamming 窗。声音信号具有准平稳特性，即信号只在短时段上才可视为是一个平稳过程。可以用平稳过程的分析方法进行分析，因此需将声音信号划分为一个个的短时段，每一短时段称为一帧，为从声音信号中切取含有 N 个样本的声音信号波形，需要用时间窗函数乘以原来的语音信号。而矩形窗的采用使得每一帧信号在起始点和终点处发生突变，导致 Gibbs 现象的出现。为了减小起始点和终点处的不连续性，用 Hamming 窗(即升余弦窗)与每一帧相乘。

3)取帧长 $N=256$ 点，对每一帧作 FFT(Fast Fourier transform)变换，对频谱取模平方得到离散功率谱。

4)把上述功率谱映射到梅尔频标，再用 M 个梅尔带通滤波器滤波，得到一组系数 $m1, m2, \cdots$. 梅尔滤波器在频域上为在梅尔频率轴上是均匀分布的重叠的三角窗。

5)将每个滤波器的输出取对数，得到相应频带的对数功率谱。

6)对梅尔对数功率谱作离散余弦变换，则得到的谱的幅值即为 MFCCs。

标准的 MFCC 只反映了语音参数的静态特性，一阶差分 MFCC(△MFCC)是一种动态参数，反映了语音参数的动态特性，有较好的鲁棒性。在一阶差分 MFCC 的基础上，还可以进一步计算出二阶差分 MFCC。

3 基于 GMM 的昆虫声音识别

3.1 混合高斯模型 GMM

混合高斯模型是单状态随机生成模型。与某一种类相关的概率密度函数是通过几个高斯函数的加权和来建模的。一个高斯混合密度是 M 个成分密度的加权和(图 1)，通过下式计算(Douglas *et al*, 1995)：

$$p(\vec{x} \mid \lambda) = \sum_{i=1}^{M} p_i b_i(\vec{x}) \tag{4}$$

其中 $\vec{x}$ 是一个 D 维的随机向量，$b_i(\vec{x})$，$i=1, \cdots, M$，是各成分的密度，p_i，$i=1, \cdots, M$，是混合权值。每个成分密度是一个 D 度可变的高斯函数，如下式：

$$b_i(\vec{x}) = \frac{1}{(2\pi)^{D/2}\left|\sum_i\right|^{1/2}} \exp\left\{-\frac{1}{2}(\vec{x} - (\vec{\mu}_i)\right\} \tag{5}$$

其均值向量为 $\vec{\mu}_i$，协方差矩阵为 $\left|\sum_i\right|$，混合权值满足关系 $\sum_{i=1}^{M} p_i = 1$。

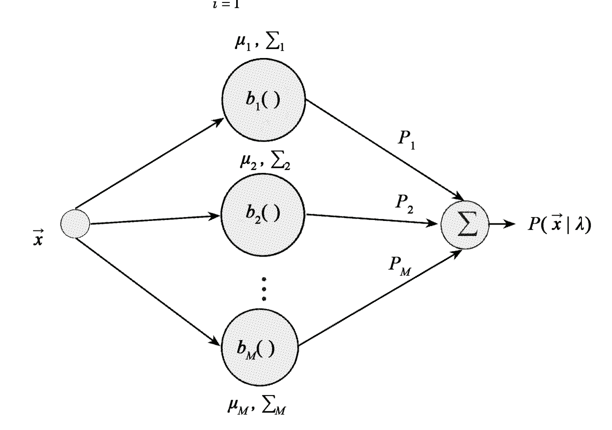


图 1 M 个成分的混合高斯密度为各成分高斯密度的加权和

Fig. 1 The Gaussian mixture density of M components is the result of weighted sum of M Gaussian densities

完整的高斯混合密度由所有成分的均值向量、协方差矩阵和混合权值参数化得到。这些参数可以和以下符号整合表示：

$$\lambda = \{p_i, \vec{\mu}_i, \sum_i\} \quad i = 1, 2, \cdots, M \tag{6}$$

每一类昆虫可以由一个 GMM 表示，并可用它的模型 λ 来指代。

3.2 期望最大化(expectation-maximuzation, EM)算法

均值向量、协方差矩阵和混合权值等 GMM 参数可以使用最大似然准则通过 EM 算法迭代估计得到。EM 算法是估计参数 λ 的经典算法，它也是最大似然估计(maximum likelihood estimation, ML)和最大后验估计(maximum a posteriori, MAP)中最常

用的方法。该方法最早由 Baum 和 Petrie（1966）及 Baum 和 Eagon（1972）提出，在隐马尔可夫模型（HMM）（Bilmes，1998）参数集的似然估计中最为常用。

EM 算法的基本思想是：从某个初始模式 λ 出发，去估计另一个新的模式 $\bar{\lambda}$，使得 $p(X \mid \bar{\lambda}) \geq p(X \mid \lambda)$；然后新的模式在下一轮迭代中作为初始模式进行估计，这个过程反复进行直到达到收敛的阈值。

在每一次 EM 迭代中，用以下重估计公式来保证模式似然值的单调递增：

混合权值：

$$\bar{p}_i = \frac{1}{T}\sum_{t=1}^{T} p(i \mid \vec{x}_t, \lambda) \tag{7}$$

均值：

$$\vec{\bar{\mu}}_i = \frac{\sum_{t=1}^{T} p(i \mid \vec{x}_t, \lambda)\vec{x}_t}{\sum_{t=1}^{T} p(i \mid \vec{x}_t, \lambda)} \tag{8}$$

协方差：

$$\bar{\sigma}_i^2 = \frac{\sum_{t=1}^{T} p(i \mid \vec{x}_t, \lambda)x_t^2}{\sum_{t=1}^{T} p(i \mid \vec{x}_t, \lambda)} - \bar{\mu}_i^2 \tag{9}$$

则声音类别 i 的后验概率计算如下：

$$p(i \mid \vec{x}_t, \lambda) = \frac{p_i b_i(\vec{x}_t)}{\sum_{k=1}^{K} p_k b_k(\vec{x}_t)} \tag{10}$$

3.3 昆虫声音识别

对于昆虫声音识别，包含 N 类昆虫的组 $S = \{1, 2, \cdots, N\}$ 可以用 N 个 GMM：$\lambda_1, \lambda_2, \cdots, \lambda_N$ 来表示。目标是对给定的观察对象找到具有最大后验概率的种类模型。可以用下列公式表示：

$$\hat{S} = \arg\max_{1 \leq k \leq N} \Pr(\lambda_k \mid X) = \arg\max_{1 \leq k \leq N} \frac{p(X \mid \lambda_k)Pr(\lambda_k)}{p(X)} \tag{11}$$

其中第 2 个等式由贝叶斯规则得到。假定所有种类相似度一致（即 $Pr(\lambda_k) = 1/N$），对所有的昆虫声音模型而言，$p(X)$ 都相同，则分类规则简化为：

$$\hat{S} = \arg\max_{1 \leq k \leq N} p(X \mid \lambda_k) \tag{12}$$

使用观察之间的独立性和计算对数，则昆虫声音的识别系统只计算：

$$\hat{S} = \arg\max_{1 \leq k \leq N} \sum_{t=1}^{T} \log p(\vec{x}_t \mid \lambda_k) \tag{13}$$

其中 $p(\vec{x}_t \mid \lambda_k)$ 由式（10）给出。

4 识别试验结果

该方法在配置为 Intel Core2 2.16GHz，1G RAM 的 PC 机上用 Matlab 实现。试验中昆虫声音有 58 种，每一种声音因录音时长不同，经预处理得到的样本数各不相同，其中一个较长的样本用于训练并建立 GMM 模型，剩下的短时样本用于测试，每一种昆虫声音用于测试的样本数介于 1 ~ 14 个之间，每一大类的总测试样本数见表 2。

表 1 给出了特征提取、训练和识别过程的时间性能，在提取 MFCCs 特征时采用了 24 组滤波器。从表中可以看出，特征提取和训练过程的时间性能跟每个样本的持续时间相关，特征提取和训练的时间与样本时长正相关，但并非线性相关，因为它们还受声音信号的功率影响；模型训练的时间还跟 GMM 的成分数有关，它与 GMM 的成分数成正比；识别时间不包括特征提取的步骤，所以其受样本时长影响相对较小。事实上，在测试阶段，使用的测试样本长度基本一致，所以表 1 中只给出一个鉴别时间值。

表 1　基于 MFCC 和 GMM 的昆虫声音自动识别算法的时间性能

Table 1　Time performance of automatic acoustic insect identification algorithm based on MFCC and GMM

声音样本持续时间（s）Duration of sound sample	特征提取时间（s）Feature extraction time	不同成分数 GMM 训练时间（s）Training time of GMMs with different number of components					鉴别时间（s）Identification time
		1	2	3	4	5	
5	0.480	0.034	0.133	0.211	0.234	1.605	0.328
15	0.735	0.109	0.404	1.727	2.194	1.900	
30	1.222	0.193	9.283	13.646	18.544	23.287	
45	7.876	0.238	9.729	21.186	28.498	37.000	

图 2 给出了 1.2 s 长的昆虫声音测试样本用不同初始化方法及不同成分数下 GMM 识别的正确率。由图可知，用模糊 C 均值聚类初始化及成分数为 2 时的 GMM 取得的识别效果最佳，其识别正确率高达 98.95%。具有较少成分的 GMM 总体上分类正确率高于较多成分的 GMM，表明昆虫声音相对人类语音而言相对单调。图 5 也表明使用模糊 C 均值聚类初始化方法总体效果要优于 k-均值聚类方法。

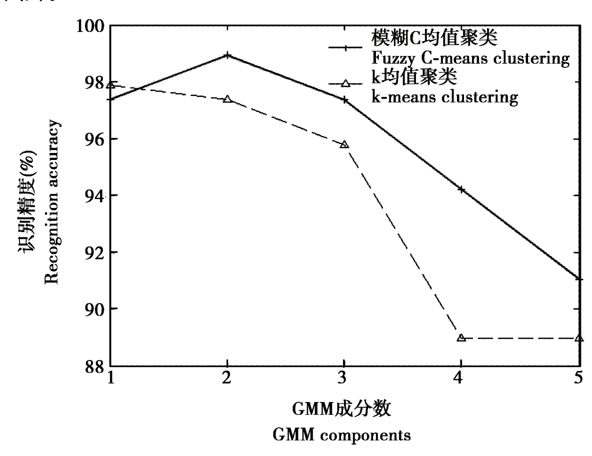


图 2 不同初始化方法及不同成分数下 GMM 昆虫声音识别精度

Fig. 2 The recognition accuracies of insect sounds with GMMs under different components and different initialization methods

表 2 给出了使用模糊 C 均值聚类初始化及成分数为 2 时的 GMM 识别各种不同昆虫声音时的最终结果。除了仓储害虫的运动和取食声的识别率相对低一点(95.65%)，其他类昆虫的总识别正确率都高于 98%。

表 2 不同类别昆虫声音的识别精度比较

Table 2 A comparison on recognition accuracy for different types of insect sounds

昆虫声音类别 Catalogues of insect sounds	识别精度 Recognition accuracy	总测试样本数(个) Overall number of testing samples
仓储害虫运动和取食声(7 类) Movement and feeding sounds of stored product insects (7 classes)	0.9565	23
土壤无脊椎动物运动和取食声 (15 类) Movement and feeding sounds of soil invertebrates (15 classes)	1.0000	48
土壤昆虫防御性摩擦声(2 类) Defensive stridulation of soil insects (2 classes)	1.0000	4
树木昆虫运动和取食声 (13 类) Movement and feeding sounds of insects in wood (13 classes)	1.0000	64
植物昆虫运动和取食声(1 类) Movement and feeding sounds of insects in plants (1 class)	1.0000	2
昆虫翅振及腹部振动声 (12 类) Vibrating sounds of wings and abdomen of insects (12 classes)	0.9804	53

5 结论

本研究使用类似于说话人识别方法的信号参数化方法和先进的模式识别技术实现了对昆虫声音的自动鉴别。提出的自动鉴定方法用 MFCC 作为声音特征、GMM 作为分类器，该方法在识别 58 类昆虫声音时取得的平均识别率为 98.95%，鉴别一个 1 s 左右的声音样本所需要的时间为 300 ms 左右，从识别正确率和识别时间上都表现出了良好的性能。然而，要将该方法付诸应用还有很多工作要做，在现实世界中，昆虫都是生活在带有背景噪声的自然环境中，这种背景噪声的存在将会影响识别算法的性能，这种背景噪声可以表现为人类活动产生的声音如飞机、汽车、机器等声音，各种自然现象带来的声音如风雨雷电等，也可能是昆虫同类活动产生的声音，在这种自然环境中使用本文算法之前首先要能去除这些背景噪声。另外，同类昆虫在不同状态下发出的声音也会有所区别，这些不同状态包括昆虫的雌雄、龄期、求偶、竞争、报警等，那么昆虫自动识别系统也应能够处理这些不同状态。本文中使用的声音文件都是从录音信号中截取了没有噪音的声音片段，在今后的研究工作中，我们将在进行识别之前，尝试从混有噪声背景的声音信号中分离并检测昆虫声音。

参考文献 (References)

Arbogast TR, Kendra EP, Mankin WR, McGovern EJ, 2000. Monitoring insect pests in retail stores by trapping and spatial analysis. *Journal of Economic Entomology*, 93(5): 1531-1542.

Baum LE, Eagon JA, 1972. An inequality and associated maximization technique in statistical estimation for probabilistic functions of Markov processes. *Inequalitie*, 3(1): 1-8.

Baum LE, Petrie T, 1966. Statistical inference for probabilistic functions of finite state Markov chains. *Ann. Math. Statistics*, 37:

1554 – 1563.

Bilmes JA, 1998. A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models. International Computer Science Institute. 1 – 13.

Brandhorst-Hubbard JL, Flanders KL, Mankin RW, Guertal EA, Crocker RL, 2001. Mapping of soil insect infestations sampled by excavation and acoustic methods. *Journal of Economic Entomology*, 94(6): 1452 – 1458.

Cheng JK, Sun YH, Ji LQ, 2010. A call-independent and automatic acoustic system for the individual recognition of animals: a novel model using four passerines. *Pattern Recognition*, 43(11): 3846 – 3852.

Chesmore ED, 2001. Application of time domain signal coding and artificial neural networks to passive acoustical identification of animals. *Applied Acoustics*, 62(12): 1359 – 1374.

Davis SB, Mermelstein P, 1980. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28(4): 357 – 366.

Ganchev T, Potamitis I, Fakotakis N, 2007. Acoustic monitoring of singing insects. In: Proceedings of the 2007 IEEE International Conference on Acoustics, Speech and Signal Processing, Honolulu. Vol. 4. 721 – 724.

Han P, 2003. Voice-pattern recognition of stored product insects. *Computer Engineering*, 29(22): 151 – 154. [韩萍, 2003. 仓储物害虫声音的模式识别. 计算机工程, 29(22): 151 – 154]

Hay-Roe MM, Mankin RW, 2004. Wing-click sounds of *Heliconius cydno alithea* (Nymphalidae: Heliconiinae) butterflies. *Journal of Insect Behavior*, 17(3): 664 – 674.

Mankin RW, 2009. Sound Library. http://www.ars.usda.gov/pandp/docs.htm?docid=10919#albopictus.

Mankin RW, Mizrach A, Hetzroni A, Levsky S, Nakache Y, Soroker V, 2008a. Temporal and spectral features of sounds of wood-boring beetle larvae: identifiable patterns of activity enable improved discrimination from background noise. *Florida Entomologist*, 91(2): 241 – 247.

Mankin RW, Osbrink WL, Oi FM, Anderson JB, 2002. Acoustic detection of termite infestations in urban trees. *Journal of Economic Entomology*, 95(5): 981 – 988.

Mankin RW, Smith MT, Tropp JM, Atkinson EB, Jong DY, 2008b. Detection of *Anoplophora glabripennis* (Coleoptera: Cerambycidae) larvae in different host trees and tissues by automated analyses of sound-impulse frequency and temporal patterns. *Journal of Economic Entomology*, 101(3): 838 – 849.

Mermelstein P, 1976. Distance measures for speech recognition, psychological and instrumental. *Pattern Recognition and Artificial Intelligence*: 374 – 388.

Pinhas J, Soroker V, Hetzroni A, Mizrach A, Teicher M, Goldberger J, 2008. Automatic acoustic detection of the red palm weevil. *Computers and Electronics in Agriculture*, 63(2): 131 – 139.

Reynolds DA, Rose RC, 1995. Robust text-independent speaker identification using Gaussian mixture speaker models. *IEEE Transactions on Speech and Audio Processing*, 3(1): 72 – 83.

Riede K, 1998. Acoustic monitoring of Orthoptera and its potential for conservation. *Journal of Insect Conservation*, 2: 217 – 223.

Thoms EM, 2000. Use of an acoustic emissions detector and intragallery injection of spinosad by pest control operators for remedial control of drywood termites (Isoptera: Kalotermitidae). *Florida Entomologist*, 83(1): 64 – 72.

Vulinec K, 2000. Dung beetles (Coleoptera: Scarabaeidae), monkeys, and conservation in Amazonia. *The Florida Entomologist*, 83(3): 229 – 241.

Wu ZJ, Cao ZG, 2005. Improved MFCC-based feature for robust speaker identification. *Tsinghua Science and Technology*, 10(2): 158 – 161.

Xue YZ, 2010. Automatic Identification of Bird Species by Their Calls and Songs. MSc Thesis, National Taipei University of Technology, Taibei, Taiwan, China. [薛宇志, 2010. 依照鸟类鸣叫声和鸣唱声识别其种类. 台北: 台北科技大学硕士学位论文]

Zhang ML, Crocker RL, Mankin RW, Flanders KL, Brandhorst-Hubbard JL, 2003. Acoustic identification and measurement of activity patterns of white grubs in soil. *Journal of Economic Entomology*, 96(6): 1704 – 1710.

Zhang WL, Liu Q, 2005. Mel-frequency cepstral coefficients extraction and its application on voiceprint recognition. *Journal of Guizhou University (Natural Sciences)*, 22(2): 208 – 210. [张万里, 刘桥, 2005. Mel 频率倒谱系数提取及其在声纹识别中的作用. 贵州大学学报(自然科学版), 22(2): 208 – 210]

Zhu LQ, Wang HB, Zhang Z, 2010. Automatic acoustical identification of insects based on MFCC and VQ. *Acta Entomologica Sinica*, 53(8): 901 – 907. [竺乐庆, 王鸿斌, 张真, 2010. 基于 Mel 倒谱系数和矢量量化的昆虫声音自动鉴别. 昆虫学报, 53(8): 901 – 907]

（责任编辑：袁德成）